

Traitement informatisé des documents épigraphiques

Joseph DENOZ et Gérard PURNELLE

I. Introduction.

Depuis le début des années 70, des spécialistes de l'épigraphie et de l'informatique appliquée aux sciences humaines ont tenté de développer des méthodes automatiques de traitement des inscriptions grecques et latines; c'est ainsi qu'en décembre 1972 déjà se tenait à Marseille une table ronde consacrée aux "applications à l'épigraphie des méthodes de l'informatique"¹.

Le Laboratoire d'Analyse Statistique des Langues Anciennes de l'Université de Liège (LASLA) qui a commencé en 1961 à enregistrer les oeuvres littéraires grecques et latines sur ordinateur, a aussi élaboré des procédures de traitement de la documentation épigraphique et papyrologique. Les résultats de ces travaux ont été diffusés notamment dans *Les Antiquités Africaines*² et dans les actes des 12^e et 13^e congrès internationaux de papyrologie³ et ont

¹ On trouvera un compte rendu de cette table ronde dans M. JANON, *Épigraphie et Informatique*, dans *Revue de l'organisation internationale pour l'étude des langues anciennes par ordinateur*, 1974, 2, pp. 1-24.

² Ét. ÉVRARD, *Une expérience de traitement automatique des inscriptions latines à l'Université de Liège*, dans *Antiquités Africaines*, 9 (1975), pp. 39-57

³ Arthur BODSON, *Exploitation automatique de la documentation papyrologique : Application à un corpus*, dans *Proceedings of the twelfth international Congress of Papyrology*, 1970, pp. 41-44; Étienne ÉVRARD, *Exploitation automatique de la documentation papyrologique : Description d'un programme d'ordinateur*, dans *Proceedings of the twelfth international Congress of Papyrology*, 1970, pp. 121-126; Alfred TOMSIN, *Les papyrologues et le travail papyrologique par ordinateur*, dans *Proceedings of the twelfth international Congress of Papyrology*, 1970, pp. 471-484 et J. DENOZ, *Applications de programmes d'ordinateur aux textes relatifs aux oasis : méthode et traitement*, dans *Actes du XIII^e Congrès international de papyrologie*, Munich, 1973, p. 415-420

abouti à la publication de trois fascicules consacrés au traitement automatique des papyrus grecs⁴.

Les progrès de l'informatique au cours des dernières années et plus particulièrement le développement de la micro-informatique nous ont conduits à élaborer une méthode nouvelle dont l'objectif est de constituer à terme une banque de données rassemblant toutes les inscriptions grecques et latines.

Une telle entreprise, cela va de soi, ne peut être menée à bien par un chercheur isolé ou par un seul laboratoire. C'est pourquoi le Centre Informatique de Philosophie et Lettres de l'Université de Liège, le Centre Pierre Paris de l'Université de Bordeaux III et le Centre d'Histoire Ancienne de l'Université de Franche-Comté (Besançon) ont décidé de travailler ensemble au Programme d'Enregistrement, de Traitement et de Recherche Automatiques en Epigraphie et de créer la Fondation européenne PETRAE qui a officiellement vu le jour en 1985 et qui est le fruit de recherches menées conjointement par les trois centres depuis le début de l'année 1984.

II. La méthode.

L'emploi de procédures automatiques pour la documentation épigraphique peut prendre plusieurs formes.

On peut se limiter à enregistrer sur support informatique le texte des inscriptions et un certain nombre d'informations historiques et techniques. Une telle documentation sera sans doute utile, mais ne conduira qu'à des exploitations qualitativement et quantitativement limitées.

De même, la préparation par ordinateur de l'édition d'un *corpus* ne conduira qu'à des applications ponctuelles dont le chercheur percevra rapidement le peu de diversité.

La conception et la réalisation d'une banque de données doivent tenir compte à la fois de la richesse et de la variété de l'information à traiter, ainsi que des possibilités d'exploitation qu'offrira la documentation. Appliqués à l'épigraphie, ces principes conduisent à accompagner chaque mot d'une inscription de renseignements nombreux et variés. La richesse des enregistrements ralentit les tâches de constitution des fichiers, mais elle s'avère indispensable si l'on veut

⁴ J. BINGEN, A. TOMSIN, A. BODSON, J. DENOZ, J.C. DUPONT, ÉT. ÉVRARD, *Choix de papyrus grecs. Essai de traitement automatique*, Liège, LASLA, 1968; A. TOMSIN, J. DENOZ, *Traitement automatique de papyrus grecs*, 2, Liège, 1971; A. TOMSIN, ÉT. ÉVRARD, J. DENOZ, P. BOLLAND, J. STRAUSS, *Traitement automatique de papyrus grecs*, 3, Liège, LASLA, 1973.

que la banque soit utile non seulement aux épigraphistes et aux historiens, mais encore aux linguistes et aux philologues.

Pour atteindre cet objectif, nous avons conçu un système de traitement des inscriptions qui se déroule en six phases successives : l'enregistrement sur micro-ordinateur du texte des inscriptions, la composition de la notice descriptive du document, que l'on appelle le lemme épigraphique, l'analyse lexicologique et morpho-syntaxique de chaque mot des inscriptions, la constitution d'un *index verborum* général, l'établissement de relevés lexicologiques et grammaticaux, la préparation d'*indices* spécialisés.

Avant de décrire en détail ces différentes opérations, on observera que le chercheur qui le souhaite peut utiliser les deux premières phases du système pour les entreprises limitées que nous avons citées précédemment.

a. l'enregistrement des inscriptions.

L'objectif de cette première phase de la procédure est de préparer un fichier dans lequel se trouvent tous les mots qui figurent dans les inscriptions traitées. L'organisation du fichier est relativement simple; chaque mot-forme constitue un enregistrement, il est accompagné d'une référence composée de trois éléments, à savoir un numéro attribué à l'inscription et, éventuellement, les indications de la colonne et du fragment. En outre, chaque mot qui figure en fin de ligne porte un code spécial qui servira à calculer automatiquement une référence très précise au moment où se préparent les indexations lexicologique et morphologique.

Nous avons choisi d'utiliser pour cette étape les micro-ordinateurs, en raison des avantages techniques et logiques que présente ce matériel.

Au point de vue technique, son coût d'utilisation est pratiquement nul, il est en tout cas infiniment moindre que celui des ordinateurs centraux. Par ailleurs, le micro-ordinateur est facilement accessible n'importe où par un chercheur qui désire enregistrer un *corpus* limité aux inscriptions qu'il étudie. En décentralisant au maximum les opérations initiales de saisie, les auteurs de PETRAE permettent à chacun de disposer de leurs logiciels.

Sur le plan de la programmation, il nous a paru indispensable que le matériel employé satisfasse à certaines exigences.

Il doit avant tout permettre de dactylographier et de faire apparaître simultanément sur l'écran différents alphabets, latin, grec ou autre. Cette caractéristique rend notamment possible le traitement d'inscriptions bilingues ou de documents qui contiennent des caractères d'origines diverses. En outre, le micro-ordinateur doit afficher des lettres pointées, ainsi que tous les signes utilisés en épigraphie.

Dans son état actuel le progiciel de saisie de PETRAE offre cinq polices de caractères, à savoir une police française, deux polices latines, la seconde étant constituée des lettres pointées, ainsi que deux alphabets grecs. La possibilité de traiter d'autres alphabets est prévue dans le système, de telle sorte que l'on pourrait employer douze jeux de caractères différents. L'utilisateur dispose aussi de signes tels que les crochets droits, les crochets brisés, les accolades, ainsi que de symboles pour noter la ligature, la césure d'un mot ou encore l'érasure.

Les conventions critiques utilisées par PETRAE respectent les principes de l'Académie de Leyde : les crochets droits marquent la restitution de lettres disparues, les crochets brisés entourent les lettres omises, les accolades signalent l'éjection des lettres interpolées et les parenthèses indiquent la résolution des abréviations. Des points entre crochets droits figurent des lettres impossibles à restituer mais dont on connaît le nombre, des tirets entourés de crochets droits signalent un nombre indéterminé de caractères perdus. Le signe + symbolise une lettre incertaine.

L'utilisation du programme d'enregistrement ne pose guère de problèmes. L'inscription est dactylographiée sous la forme que l'épigraphiste lui a donnée après analyse critique : toutes les lettres disparues ou omises sont restituées, toutes les abréviations sont résolues, les lettres douteuses sont pointées et, dans les inscriptions grecques, les accents et les esprits sont indiqués.

Quand le texte de l'inscription a été dactylographié sous sa forme développée, le programme interprète les signes critiques et génère deux états différents du document. La première recomposition présente, en lettres capitales, le texte tel qu'il est gravé sur le support, tandis que la seconde fait apparaître un état reconstitué en caractères majuscules et minuscules et, pour le grec, avec les esprits et les accents. Ainsi, si l'on dactylographie *[Se]quan(us) viz(it) annos*, le programme présentera l'état majuscule *[- -]QVAN.VIX.ANNOS* et, dans ce cas simple, un état minuscule identique au texte saisi, soit *[Se]quan(us) viz(it) annos* (cf. document 1).

Le double affichage permet de vérifier les informations enregistrées et de les comparer au document qui porte le texte de l'inscription. A ce stade, il est possible de modifier aisément les erreurs repérées dans la mesure où le logiciel intègre des procédures sophistiquées de correction.

Le programme de saisie est accompagné d'un module d'impression qui présente à son tour les deux états du texte. Ce module est disponible soit sur une imprimante matricielle (cf. document 1), soit sur une imprimante à laser dont la qualité d'impression est supérieure (cf. document 2).

INSCRIPTION 000003 0000

ETAT DE LA PIERRE

- 1 I.Q.M
- 2 SAÇ[...]M.ET
- 3 ΣΑΡΑΤ[....]Τ
- 4 ΑΡΤ[...]ΙΔΙ

ETAT APRES DEVELOPPEMENT

- 1 I(oui) O(ptimo) M(aximo)
- 2 sac(ru)m et
- 3 Σαράπιδι ε]t
- 4 Ἀρτ[έμ]ιδι

Imprimante FX/BO - C.I.P.L.

Document 1.

INSCRIPTION 000025 0000

Etat de la pierre

- 1 I.Ο.Μ
- 2 SACR[...]
- 3 ET.ΣΑΡΑΠ[...]

Etat après développement

- 1 I(oui) Ο(ptimo) M(aximo)
- 2 sacr[um]
- 3 et Σαράπ[ιδι]

INSCRIPTION 000064 0000

Etat de la pierre

- 1 IN.H.D.[...]
- 2 APOL.V[...]
- 3 OROL.P[...]

Etat après développement

- 1 in h(onorem) d(omus) [d(iuinae) D(eo)]
- 2 Apol(lini) u[ic(ani)]
- 3 Orol(aunenses) p[os(uerunt)]

Imprimante QMS SmartWriter - C.I.P.L.

Document 2.

b. composition de la notice descriptive du document.

Après avoir enregistré une série d'inscriptions, il convient de constituer pour chacune d'entre elles une notice descriptive qui sera intégrée à la banque de données textuelles.

A priori, il semble difficile de rassembler dans un ensemble logiquement construit toutes les informations qui font partie du "lemme épigraphique" parmi lesquelles doivent figurer de manière relativement détaillée les principaux renseignements techniques, historiques, géographiques, linguistiques et bibliographiques relatifs à l'inscription sur laquelle porte la notice.

Les informations se présentent sous diverses formes, elles sont de longueur fixe lorsqu'il s'agit de données numériques (date de l'inscription, dimensions du support) ou de données codifiées (langue) mais il s'agit plus souvent d'informations de longueurs variables. Parmi celles-ci, il faut distinguer celles qui doivent s'intégrer à une liste fermée de descripteurs (types de support, régions géographiques) et celles qui se présentent comme du "texte libre" (commentaires, ...).

Les progiciels commerciaux tels que DBASE, par exemple, ne permettent pas de traiter avec toute la souplesse souhaitée des informations aussi complexes que celles que nous rencontrons dans la plupart des applications de l'informatique aux sciences humaines; c'est la raison pour laquelle nous utilisons pour ce fichier documentaire le logiciel PALAMÈDE qui a été spécialement conçu pour la saisie d'informations structurées⁵. Le principal avantage de PALAMÈDE est qu'il permet de définir le type d'enregistrements que l'on traite, ainsi que les rubriques et leurs modalités de saisie, quels que soient leur nombre et leur nature.

c. Analyses lexicographique et grammaticale.

Le texte des inscriptions enregistrées au moyen du programme décrit précédemment (*sub a*) est conservé sur disquette dans un fichier qui, transféré sur gros ordinateur, sera utilisé pour préparer l'indexation.

La Fondation PETRAE bénéficie dans ce domaine de l'expérience acquise par le LASLA dont elle utilise les procédures semi-automatiques de lemmatisation et d'analyse des langues grecque et latine.

Sans entrer dans le détail, rappelons que les logiciels de Liège se fondent sur l'exploitation d'un lexique latin et d'un lexique grec de plusieurs milliers

⁵ Ce logiciel a fait l'objet d'une description détaillée dans : G. PURNELLE, *PALAMÈDE, Logiciel de saisie d'informations structurées sur micro-ordinateur*, dans *Revue, Informatique et statistique dans les sciences humaines*, 22 (1986), pp. 157-169.

de mots et, pour la langue latine, d'une liste complète des désinences. Grâce à ces fichiers, les programmes produisent automatiquement pour chaque mot rencontré dans un texte, tous les lemmes et toutes les analyses grammaticales formellement possibles (cf. document 3).

En latin, une analyse se compose en moyenne de cinq ou six informations codées explicitant l'analyse morphologique. Ainsi, pour la forme *templo*, l'ordinateur propose le lemme *templum* et deux analyses codées 12E et 12F ce qui signifie respectivement substantif de la deuxième déclinaison au datif singulier et substantif de la deuxième déclinaison à l'ablatif singulier. Le mot *nonebat* sera lemmatisé *noneo* et accompagné de l'analyse 52C12, c'est-à-dire verbe (5) de la deuxième conjugaison à la voix active (2), troisième personne du singulier (C), indicatif (1) imparfait (2).

Lorsqu'une forme reçoit plusieurs analyses, il incombe au chercheur de sélectionner celle qui est pertinente en fonction du contexte et de rejeter les autres. Le philologue devra également indiquer le lemme et l'analyse des formes que l'ordinateur n'a pu identifier⁶.

À la suite des opérations de lemmatisation et d'analyse, le chercheur possède un fichier où chaque forme du texte est rapportée à un seul lemme et est affectée d'une analyse correcte en fonction du contexte. Ce nouveau fichier est conservé sur un support informatique (disque ou bande) et va dès ce moment enrichir la banque de données et se prêter à diverses exploitations : celle-ci fournira des listes exhaustives de faits de langue aux grammairiens, aux philologues et aux linguistes; les épigraphistes y chercheront les noms de choses ou de personnes ou encore les formules qui les intéressent. La masse des informations ainsi conservées et l'usage qui est fait de l'ordinateur pour l'interroger sont particulièrement bien adaptés non seulement aux études quantitatives mais aussi à toutes espèces de recherches qualitatives.

d. l'index verborum général.

À partir du fichier dans lequel se trouvent les mots lemmatisés et analysés, l'ordinateur produit par une opération simple de classement logique, un *index verborum* complet du *corpus* traité. Cette étape, entièrement automatique, est réalisée en un temps très bref grâce aux logiciels que le LASLA a élaborés pour publier ses propres *indices* d'auteurs latins et grecs.

⁶ Au sujet des programmes de lemmatisation et d'analyse on consultera, pour la langue latine J. DENOZ, *Recherches sur le traitement automatique de la langue latine*, dans *Revue de l'Organisation internationale pour l'étude des langues anciennes par ordinateur*, 1973, 1, pp. 1-89 et pour le grec J. DENOZ, *Aristote, Poetica, Index verborum, Liste de fréquence*, Liège, CIPL-LASLA, 1988, pp. I à XII.

NE	NE	1	60000	0001	01
	NE	2	670	0001	02
	NE	3	68000	0001	03
	NE	4	82032	0001	04
FASTIDIOSUM	FASTIDIOSVS		21A00 6	0002	01
	FASTIDIOSVS		21C00 5	0002	02
ESSET	EDO	2	56C32	0003	01
	SVM	1	56C32- DD	0003	02
	SVM	2	E6C32	0003	03
CLEMENTIAE	CLEMENTIA		11D00	0004	01
	CLEMENTIA		11E00	0004	02
	CLEMENTIA		11J00	0004	03
TUAE	TVVS		42D00 2	0005	01
	TVVS		42E00 2	0005	02
	TVVS		42J00 2	0005	03
CONSTANTINE				0006	01
MAXIME	MAGNVS		21B00 4	0007	01
SINGULOS	SINGVLVS		33L00 4	0008	01
QUOSQUE	QVISQVE	1	46L 4	0009	01
	QVISQVE	2	48L00 4	0009	02
PRINCIPES	PRINCEPS	1	13J00	0010	01
	PRINCEPS	1	13L00	0010	02
	PRINCEPS	2	22J00 3	0010	03
	PRINCEPS	2	22L00 3	0010	04
UEL	VEL	1	60000	0011	01
	VEL	2	81000	0011	02
PRINCIPUM	PRINCEPS	1	13M00	0012	01
	PRINCEPS	2	13M00	0012	02
LIBEROS	LIBER	N	12L00	0013	01
	LIBER	2	21L00 4	0013	02
	LIBERI		12L00	0013	03
PER	PER		70300	0014	01
LIBROS	LIBER	1	12L00	0015	01
SINGULOS	SINGVLVS		33L00 4	0016	01
LEGERE	LEGO	2	53L14	0017	01
	LEGO	2	53071	0017	02
ADHIBUI	ADHIBEO		52A14&	0018	01
MODERATIONEM	MODERATIO		13C00	0019	01

Document 3.

Dans l'*index* d'un texte grec, les mots sont classés en ordre alphabétique et les occurrences d'un même mot apparaissent à la suite du lemme dans l'ordre alphabétique des formes; chacune de celles-ci est accompagnée de sa référence. Lorsqu'une même forme est employée plusieurs fois dans un même *corpus*, ses références sont rangées en ordre croissant. Le document 4 présente une page de l'*index* de la *Poétique* d'Aristote.

Pour le latin, le classement diffère de ce qui précède. Sans doute, ici aussi les lemmes sont en ordre alphabétique, mais sous chaque lemme les formes sont rangées en fonction de l'analyse grammaticale (par exemple, les singuliers avant les pluriels, le nominatif avant le vocatif, ...). Nous reproduisons au document 5 un échantillon de l'*index verborum* de l'*Art d'aimer*⁷.

e. les relevés lexicologiques et grammaticaux.

A partir du fichier indexé, l'ordinateur fournit, si on le désire, différents relevés, listes de fréquence des lemmes ou tableaux grammaticaux.

Deux listes peuvent être imprimées : l'une présente les mots rangés en fonction de leur fréquence croissante ou décroissante dans le *corpus* (cf. document 6), la seconde donne le nombre d'emplois de chaque mot dans un *corpus* déterminé et dans les sous-ensembles de ce *corpus* définis par le chercheur (cf. document 7).

Cette deuxième liste se révélera particulièrement utile en épigraphie pour étudier la dispersion de certains faits; elle rendra plus aisées et plus rigoureuses des recherches sur la diffusion du culte de telle divinité, sur le rôle d'un personnage dans plusieurs régions, ou, d'un point de vue diachronique, sur la permanence de certaines familles à l'exercice du pouvoir.

Les tableaux grammaticaux fournissent des informations qui intéressent sans doute les philologues et les linguistes mais qui seront également utiles à l'historien soucieux d'étudier les habitudes des graveurs.

f. les indices spécialisés.

Il s'agit des *indices* particuliers (*nominum, rerum*, etc.) qui se trouvent en général dans les éditions de *corpora* épigraphiques. La réalisation de tels *indices* à partir des fichiers informatisés peut elle aussi, pour une grande part, se faire automatiquement.

⁷ Cf. PURNELLE-SIMART, G. PURNELLE, *OVIDE, Ars amatoria, Remedia amoris, De medicamine, Index verborum, Listes de fréquence, Relevés grammaticaux*, Liège, CIPL-LASLA, 1987.

1447 a 23	1447 a 12	6 ἄτεχνος
ἄρμον(αν	ἄρξαμένοις	ἄτεχνοι
1449 b 29	1448 b 29	1454 b 31
ἄρμονίας	ἄρξαντος	ἄτεχνοτάτη
1447 a 26	1456 a 30	1454 b 20
1448 b 20	ἄρχεσθαι	ἄτεχνότατον
1449 a 28	1450 b 34	1450 b 18
3 ἄρσην	ἦρξεν	ἄτεχνότεραι
ἄρρενα	1449 b 7	1454 b 29
1458 a 9	1 ἄρχων	ἄτεχνότερον
1458 a 10	ἄρχων	1453 b 8
1458 a 14	1449 b 2	1 ἀτιμάζω
2 ἄρβω	6 ἥσημος	ἀτιμαζομένου
ἄρύσαι	ἥσημος	1448 a 38
1457 b 15	1456 b 35	1 ἥτιμος
1457 b 15	1456 b 38	ἥτιμοτάτων
"ἄρύσας"	1457 a 4	1448 b 12
1457 b 14	1457 a 6	4 ἥτοπος
3 ἀρχαῖος	ἥσημου	ἥτοπον
ἀρχαῖοι	1457 a 33	1460 a 35
1450 b 7	1457 a 34	1460 b 2
ἀρχαῖος	[ἥσημος]	1461 b 5
1458 b 7	1457 a 8	ἥτοπώτερον
ἀρχαίων	1 ἁσθένεια	1460 a 2
1459 b 20	ἁσθένειαν	1 ἀτράγωδος
16 ἀρχή	1453 a 34	ἀτραγυδοτάτον
ἀρχάς	0 ἁσθενικός	1452 b 37
1448 a 25	"ἁσθενικός"	1 ἀτυχέω
ἀρχή	1458 b 27	ἀτυχεῖν
1450 a 38	3 ἁσπίς	1452 b 2
1450 b 28	ἁσπίδα	2 ἀτυχία
ἀρχῆ	1457 b 22	ἀτυχίαν
1457 a 3	1457 b 32	1455 b 28
ἀρχήν	ἁσπίς	ἀτυχίας
1450 b 27	1457 b 21	1452 b 37
1457 a 6	"ἁσπίδα"	1 αὐ
1459 a 20	1457 b 21	αὐ
1459 a 32	1 ἁστήρ	1453 a 1
1459 b 19	ἁστέρας	1 ἀβλέω
ἀρχῆς	1454 b 22	αὐλᾶσιν
1448 b 22	1 ἁστυ	1461 b 32
1449 a 9	ἁστεως	1 ἀβλησις
1449 b 1	1448 a 38	αὐλήσει
1449 b 6	1 ἁστυδάμας	1448 a 9
1455 b 27	ἁστυδάμαντος	1 ἀβλητής
1455 b 29	1453 b 33	αὐληταί
1460 a 33	1 ἀσύμμετρος	1461 b 31
1 ἀρχιτεκτονικός	ἀσύμμετρον	2 ἀβλητικός
ἀρχιτεκτονικήν	1461 a 13	αὐλητική
1458 b 11	0 ἀτειρής	1447 a 24
6 ἄρχω	"ἀτειρέι"	αὐλητική
ἀρξάμενοι	1457 b 14	1447 a 15

Document 4.

<i>Rem.</i> , 63	<i>Ars</i> 1, 17	<i>ancillula</i>
<i>Rem.</i> , 541	<i>Ars</i> 1, 30	<i>Rem.</i> , 639
<i>Rem.</i> , 559	<i>Ars</i> 2, 497	2 ANCORA
<i>Rem.</i> , 657	<i>Ars</i> 3, 559	<i>ancora</i>
<i>Rem.</i> , 733	<i>Amori</i>	<i>Ars</i> 1, 772
<i>Rem.</i> , 749	<i>Ars</i> 1, 7	<i>Rem.</i> , 447
<i>amoris</i>	<i>Ars</i> 1, 79	6 ANDROMACHE N
<i>Ars</i> 2, 16	<i>Ars</i> 1, 83	<i>Andromache</i>
<i>Ars</i> 2, 176	<i>Amore</i>	<i>Ars</i> 2, 645
<i>Ars</i> 2, 242	<i>Ars</i> 1, 236	<i>Ars</i> 3, 109
<i>Ars</i> 2, 248	<i>Amores</i>	<i>Andromache</i>
<i>Ars</i> 3, 441	<i>Rem.</i> , 379	<i>Ars</i> 3, 519
<i>Rem.</i> , 143	<i>Amorum</i>	<i>Andromaches</i>
<i>Rem.</i> , 643	<i>Ars</i> 3, 343	<i>Rem.</i> , 383
<i>amori</i>	1 AMPHILOCHVS N	<i>Andromache</i>
<i>Ars</i> 1, 49	<i>Amphilochi</i>	<i>Ars</i> 2, 709
<i>Ars</i> 1, 163	<i>Rem.</i> , 455	<i>Rem.</i> , 384
<i>Rem.</i> , 759	3 AMPLEXVS	3 ANDROMEDA N
<i>amore</i>	<i>amplexus</i>	<i>Andromedan</i>
<i>Ars</i> 1, 283	<i>Ars</i> 1, 770	<i>Ars</i> 1, 53
<i>Ars</i> 1, 327	<i>Ars</i> 3, 732	<i>Andromedae</i>
<i>Ars</i> 1, 372	<i>Rem.</i> , 668	<i>Ars</i> 2, 643
<i>Ars</i> 1, 736	1 AMYCLAE N	<i>Ars</i> 3, 429
<i>Ars</i> 2, 124	<i>Amyclis</i>	1 ANGO
<i>Ars</i> 2, 519	<i>Ars</i> 2, 5	<i>angat</i>
<i>Ars</i> 2, 563	1 AMYCLAEVS N	<i>Rem.</i> , 571
<i>Ars</i> 2, 684	<i>Amyclaeis</i>	4 ANGVIS
<i>Ars</i> 3, 448	<i>Rem.</i> , 707	<i>angues</i>
<i>Ars</i> 3, 571	1 AMYGDAVM	<i>Med.</i> , 39
<i>Ars</i> 3, 682	<i>amygdala</i>	<i>angues</i>
<i>Rem.</i> , 21	<i>Ars</i> 3, 183	<i>Ars</i> 1, 187
<i>Rem.</i> , 776	1 AMYNTORIDES N	<i>anguis</i>
<i>amores</i>	<i>Amyntorides</i>	<i>Ars</i> 3, 7
<i>Ars</i> 3, 27	<i>Ars</i> 1, 337	<i>anguibus</i>
<i>amores</i>	7 AN	<i>Ars</i> 3, 77
<i>Ars</i> 2, 385	<i>an</i>	1 ANGVSTVS
<i>Ars</i> 2, 639	<i>Ars</i> 1, 375	<i>angustum</i>
<i>Rem.</i> , 557	<i>Ars</i> 1, 739	<i>Ars</i> 3, 274
<i>Med.</i> , 27(5)	<i>Ars</i> 2, 715	2 ANHELITVS
31 AMOR N	<i>Ars</i> 3, 472	<i>anhelitus</i>
<i>Amor</i>	<i>Rem.</i> , 556	<i>Ars</i> 1, 521
<i>Ars</i> 1, 4	<i>Rem.</i> , 797	<i>Ars</i> 3, 803
<i>Ars</i> 1, 21	<i>Rem.</i> , 798	1 ANHELO
<i>Ars</i> 1, 23	1 ANALEMPTRIS	<i>anhelat</i>
<i>Ars</i> 1, 232	<i>analemptrides</i>	<i>Rem.</i> , 394
<i>Ars</i> 2, 17	<i>Ars</i> 3, 273	46 ANIMVS
<i>Ars</i> 2, 158	9 ANCILLA	<i>animus</i>
<i>Ars</i> 2, 229	<i>ancilla</i>	<i>Ars</i> 2, 535
<i>Ars</i> 2, 708	<i>Ars</i> 1, 367	<i>Ars</i> 3, 467
<i>Ars</i> 3, 436	<i>Ars</i> 2, 525	<i>Ars</i> 3, 707
<i>Rem.</i> , 1	<i>Ars</i> 3, 607	<i>Ars</i> 3, 763
<i>Rem.</i> , 39	<i>Ars</i> 3, 665	<i>Rem.</i> , 123
<i>Rem.</i> , 148	<i>ancillam</i>	<i>animum</i>
<i>Rem.</i> , 198	<i>Ars</i> 1, 351	<i>Ars</i> 1, 115
<i>Rem.</i> , 246	<i>ancillae</i>	<i>Ars</i> 1, 619
<i>Rem.</i> , 346	<i>Ars</i> 3, 485	<i>Ars</i> 1, 733
<i>Rem.</i> , 358	<i>ancillae</i>	<i>Ars</i> 2, 119
<i>Rem.</i> , 530	<i>Ars</i> 2, 257	<i>Ars</i> 2, 250
<i>Rem.</i> , 551	<i>ancilla</i>	<i>Ars</i> 3, 499
<i>Rem.</i> , 612	<i>Ars</i> 1, 386	<i>Rem.</i> , 129
<i>Rem.</i> , 662	<i>ancillas</i>	<i>Rem.</i> , 197
<i>Amoris</i>	<i>Ars</i> 2, 251	<i>Rem.</i> , 559
<i>Ars</i> 1, 8	1 ANCILLVLA	<i>Rem.</i> , 805

Document 5.

4	λοιπός	3	διορίζω	3	προσβολή
4	μακρός	3	δύναμαι	3	πρόσειμι 1
4	μείγνυμι	3	εϊκών	3	πρόσωπον
4	μετάσσεις	3	ἐκάτερος	3	πρότερος
4	μεταξύ	3	ἐκεῖ	3	πτῶσις
4	μηδέ	3	ἔμπεδοκλῆς	3	ῥῆσις
4	μήτηρ	3	ἐνταῦθα	3	Σικελία
4	ὄθεν	3	ἐπεισοδιόω	3	στάσιμος
4	ὄλιγος	3	ἐπεικῆς	3	στοιχεῖον
4	πολλάκις	3	ἐπισκοπέω	3	σχεδόν
4	πόσος	3	ἐπιτίθημι	3	σῶμα
4	πρό	3	ἐπιτίμημα	3	ταπεινός
4	προστίθημι	3	ἐποποιῖκός	3	τελευτή
4	Σ	3	εὐθύς	3	τετράμετρος
4	σαφής	3	ζεῦξις	3	τύχη
4	στοχάζομαι	3	ἡδύνω	3	ὑπεναντίος
4	συλλαβή	3	ἡδύς	3	ὑπερβάλλω
4	συλλογίζομαι	3	ἥλιος	3	ὑποκριτικός
4	σύνδεσμος	3	θάνατος	3	ὑπολαμβάνω
4	τέλειος	3	θέατρον	3	ὑστερος
4	ὑποτίθημι	3	θυέστης	3	φῆρω
4	ὑστερον	3	ἱκανός	3	φιάλη
4	χωρίς	3	ἴνα	3	φιλάνθρωπος
3	Λ	3	κακία	3	φίλια
3	ἄγνοια	3	κακῶς	3	φορτικός
3	ἄγών	3	κεῖμαι	3	χαίρω
3	ἀδελφή	3	κεράννυμι	3	χορικός
3	αἰδῶ	3	κίνησις	3	χρή
3	αἵνιγμα	3	κοινωνέω	2	ἄγω
3	αἰσχυρός	3	κρείσσω	2	ἀδιαίρετος
3	ἀκουστός	3	κωλύω	2	δήθης
3	ἄλκμαίων	3	μέντοι	2	Ἀθηναῖος
3	ἀμάρτημα	3	μεταβολή	2	αἰσθάνομαι
3	ἀνδρεῖος	3	μιαρός	2	αἵσθησις
3	ἀνὴρ	3	μιμητής	2	ἀκριβός
3	ἀνώμαλος	3	μοχηρία	2	ἀλλότριος
3	ἀπαγγέλλω	3	νόμος	2	ἀναγκάζω
3	ἀποδείκνυμι	3	ὀλιγάκις	2	ἀναγνωρισμός
3	ἀποδίδωμι	3	ὀνομάζω	2	ἀναιρέω
3	ἀρετή	3	ὅπως	2	ἀνάξιος
3	ἔρσην	3	παθητικός	2	ἀντέχω
3	ἀρχαῖος	3	πάλιν	2	ἀνώνυμος
3	ἀσπίς	3	παραλογισμός	2	ἄξιος
3	βαρβαρισμός	3	πάροδος	2	ἀπαγγεῖα
3	γράφω	3	περαίνω	2	ἀπεργασία
3	δηλός	3	πιθανός	2	ἀπίθανος
3	διαίρεω	3	πολύγνωτος	2	ἀπολαμβάνω
3	διάλεκτος	3	πονηρία	2	ἀπόπλοος 1
3	δίδωμι	3	ποτε	2	ἀποφαίνω
3	διηγηματικός	3	πρόβλημα	2	ἀπρεπής
3	διθύραμβος	3	πρόλογος	2	ἄρης
3	διόπερ	3	προσαγορεύω	2	ἄρθρον

Document 6.

	Total	Ars 1	Ars 2	Ars 3	Med.	Rem.
PROTEGO	1	0	1	0	0	0
PROTERVVS	2	1	0	0	0	1
PROTEVS N	1	1	0	0	0	0
PROTINVS	6	3	0	0	1	2
PROVEHO	1	0	0	1	0	0
PROVENIO	1	0	0	1	0	0
PROXIMITAS	1	0	1	0	0	0
PRVDENTIA	1	0	1	0	0	0
PVBES 1	2	1	1	0	0	0
PVBLICVS 2	4	2	0	1	0	1
PVDENDVS	6	0	2	3	0	1
PVDENS	1	1	0	0	0	0
PVDEO	8	2	0	4	0	2
PVDICVS	1	1	0	0	0	0
PVDOR	19	5	8	4	0	2
PVDOR N	1	1	0	0	0	0
PVELLA	111	34	21	29	3	24
PVER	30	8	9	8	0	7
PVERILIS	1	0	0	0	0	1
PVGNO	10	4	1	1	0	4
PVLCHER	3	1	1	1	0	0
PVLLVM 2	2	0	0	2	0	0
PVLLVS 2	1	0	0	1	0	0
PVLPITVM	1	1	0	0	0	0
PVLSO	4	2	0	2	0	0
PVLVEREVS	1	0	0	0	1	0
PVLVINVS	1	1	0	0	0	0
PVLVIS	4	2	0	1	1	0
PVMEX	1	1	0	0	0	0
PVPPIS	4	2	0	0	0	2
PVRPVRA	1	0	0	0	0	1
PVRPVREVS	7	2	1	3	0	1
PVRVS	4	0	0	3	1	0
PVSILLVS	1	0	0	0	0	1
PVTEAL N	1	0	0	0	0	1
PVTO	33	7	10	8	0	8
PYLADES N	2	1	0	0	0	1
PYRETHRVM	1	0	1	0	0	0
PYXIS	2	0	0	1	0	1
QUAEQUE	0	0	0	0	0	0
QVA 1	19	7	4	4	0	4
QVADRPES 1	2	2	0	0	0	0
QVAERO	33	10	5	8	1	9
QVALIS 2	1	0	0	0	0	1
QVALISCVMOVE 2	1	0	1	0	0	0
QVAM 1	24	3	4	7	2	8
QVAM 2	11	2	2	2	0	5
QVAMLIBET 2	3	0	0	3	0	0
QVAMVIS 2	14	2	2	3	2	5
QVAMVIS 3	1	0	0	0	0	1
QVANDO 4	1	0	1	0	0	0
QVANTO 1	1	0	0	0	0	1
QVANTVLVSCVMOVE	1	0	0	1	0	0
QVANTVM 1	1	0	1	0	0	0
QVANTVM 2	1	1	0	0	0	0
QVANTVM 3	4	1	0	1	1	1
QVANTVM 4	3	0	0	2	0	1
QVANTVS 1	3	0	2	0	1	0
QVANTVS 2	3	2	1	0	0	0
QVARE 2	3	0	1	0	0	2

Document 7.

Ainsi, dans la mesure où le système de lemmatisation développé par le LASLA a été adapté aux particularités de l'épigraphie, on établira sans difficulté un *index* des tribus, des divinités ou encore des lieux géographiques. En effet, les lemmes des noms propres sont affectés d'un code qui permet de les regrouper.

On pourra aussi isoler, éventuellement avec leur contexte, les mots appartenant à un même champ sémantique ou thématique : il suffira de les repérer dans l'*index verborum* ou dans la liste de fréquence, puis de demander à l'ordinateur de les extraire du fichier général.

Les fichiers, les *indices* et les relevés que nous venons de décrire serviront aussi à l'édition de recueils d'inscriptions et épargneront au chercheur la tâche fastidieuse qu'est la correction des épreuves. En effet, une imprimante à laser (telle que celle qui nous sert à imprimer cette revue) fournit des documents qu'il suffit de reproduire en offset.

Comme nous venons de le voir, PETRAE est à même d'éditer un *corpus* d'inscriptions et il peut en réaliser les *indices* philologiques ou épigraphiques et les imprimer. Mais son objectif essentiel reste la constitution d'une banque de données épigraphiques.

III. Quelques travaux en cours de réalisation.

Depuis sa fondation officielle en 1986, la Fondation PETRAE a commencé à enregistrer des *corpora* d'inscriptions et elle ne cesse d'accueillir de nouveaux chercheurs. A l'heure actuelle, sont en cours de traitement, soit au Centre Pierre Paris de Bordeaux, soit à l'Université de Liège les inscriptions des régions suivantes : Doride, Mytilène, Lydie, Carie-Ionie, Pont Gauche, Macédoine, Egypte, Aquitaine, Belgique, Bretagne et Catalogne. A ces régions s'ajoutent des dossiers thématiques sur le culte de Jupiter dans la péninsule Ibérique et sur les inscriptions relatives aux soldats gaulois. Enfin, un chercheur des Facultés Notre-Dame (Namur-Belgique) a entrepris de traiter tous les diplômes militaires romains.

IV. Une banque de données des langues grecque et latine.

Bien que les oeuvres littéraires et les documents épigraphiques que l'Antiquité gréco-romaine nous a légués constituent un ensemble clos et fini, il est impossible pour un chercheur d'appréhender une matière aussi étendue par des moyens traditionnels.

Or pour les disciplines littéraires et linguistiques, c'est dans le domaine

des recherches à grande échelle qu'il reste peut-être le plus à faire. Il est dès lors primordial de mettre sur pied de vastes banques de données regroupant la production d'une langue.

C'est dans cette perspective que PETRAE se propose de réunir en un fichier unique toute la documentation épigraphique. En ce qui concerne la littérature, la Fondation aura la possibilité d'utiliser et de consulter la banque de données du LASLA, profitant ainsi de l'expérience et des richesses documentaires de ce laboratoire.

Cette collaboration autorisera dans l'avenir des études de masse portant sur l'histoire de l'Antiquité et sur les deux langues classiques. Elle permettra ainsi de disposer d'informations qui se complètent mutuellement alors que à l'heure actuelle le chercheur est souvent contraint de se limiter à la seule littérature ou à un *corpus* unique d'inscriptions.

Le fait que la banque de données générale soit centralisée n'empêchera nullement la diffusion partielle de fichiers restreints. En effet, PETRAE mettra à la disposition des chercheurs qui en feront la demande des disquettes contenant des *corpora* thématiques ou géographiques d'inscriptions.

V. En guise de conclusion.

On jugera peut-être ambitieux, voire utopiques les objectifs de la Fondation européenne PETRAE. Sans doute sommes-nous encore fort éloignés de notre but, peut-être faudra-t-il attendre le siècle prochain avant d'avoir rassemblé toutes les inscriptions de l'Antiquité gréco-romaine, mais nous voulons espérer que l'enrichissement progressif de nos fichiers, la participation occasionnelle de chercheurs ou la collaboration permanente d'équipes contribueront à développer notre projet et rendront possibles de nouvelles recherches.

Pour terminer nous voudrions exprimer nos remerciements aux gouvernements français et belge qui apportent leur aide à la Fondation PETRAE par le biais des Relations culturelles internationales. En outre, PETRAE a reçu dès 1986, le soutien du Fonds National de la Recherche Scientifique de Belgique grâce auquel nous avons pu notamment acquérir une partie du matériel électronique nécessaire à nos travaux.